A cellular automata model for the study of small-size urban areas

Centre de Política de Sòl i Valoracions
Barcelona
23 January 2007

Nuno Norte Pinto
School of Technology and Management
Polytechnic Institute of Leiria
Portugal

Context and Motivation

Context [1]

- A rapid growth of the Portuguese urban areas was registered in the past two decades
- Important changes in the national planning legal framework took place in the early 1990s
- By the late 1980s a great majority of the Portuguese municipalities did not had any type of master plan. Only in 1995 all the municipalities had their municipal master plans legally approved
- Significant evolution occurred in small and peripheral urban areas, from an unsatisfactory situation in terms of urban quality – public spaces and facilities, infrastructures – to a satisfactory level of urban quality

Context and Motivation

Context [2]

- Small urban areas were poorly served by public facilities and infrastructures in the early 1990s. They lacked urban functions and, more generally, urban quality
- Their potential for attracting new populations was very low and generally they have lost population in a constant decay since the 1950s
- As a result of the use of the European Union's structural funds, these small urban areas made an important effort during the past 15 years to build new public facilities, network infrastructures, and public areas
- Some of these small municipalities presented the highest growth rates for population and constructed urban areas during the last decade

Context and Motivation

Motivation [1]

- The use of modelling as an effective planning tool was not (and still is not) a common practice among the Portuguese planning practitioners
- The administration still has the need for more planning technicians that could integrate traditional planning approaches with modern modelling techniques
- Modelling is believed to be a valid and yet unexplored tool for understanding the Portuguese urban phenomena
- It is also challenging to develop a modelling approach that could be able to simulate land use change in a context of rapid growth, based only in recent data

Context and Motivation

Motivation [2]

- The research on CA and urban studies has increased both in theoretical and operational perspectives all over the world. Many issues regarding CA and its applications to geography are still under intensive research
- The application of CA models, usually to large metropolitan areas, is based on the use of remote sensing maps and regular cell spaces. This approach is thought to be unfitted for the common Portuguese urban area
- The use of irregular cells based on census tracts is assessed, as well as the application of CA models to problems focused on small urban areas

Methodology

The use of Cellular Automata [1]

- Two main issues were considered at early stages of model design:
  - How could the growth of small urban areas be studied following an approach based on urban modelling?
  - How could the model be tailored to reproduce problems that are intrinsically simple, no matter what modelling approach would be chosen?
- Small urban areas tend to have less land uses, thus less urban functions than large ones
- The modelling technique chosen should be able to replicate this simple set of land use interactions

Methodology

The use of Cellular Automata [2]

Why the choice over CA instead of other microsimulation techniques (such as agent-base simulation, ABS) that are also capable of reproducing spatiality for different types of problems?

Two main characteristics

Spatiality

- Both CA and ABS are able to simulate well spatial relations based on sets of simple relations

Data Availability

- CA can use location-based data, as it simulates interactions between locations (and its land uses)
- ABS simulates individual behaviours (households, land parcels, trips), demanding disaggregate information that is harder to collect and to validate

Methodology

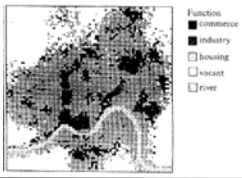
Cellular Automata structure

- CA are a very intuitive and straightforward technique and they are easily applied to spatial problems of any kind
- The technique is based on a set of rules that operate over a spatial (cellular) structure thus reproducing spatial phenomena that evolves over space. It also simulates the evolution over time
- CA have five major components
  - Cells
  - Cell space
  - Cell neighbourhood
  - Transition rules
  - Time

Methodology

Cells and Cell Space [1]

- The structural base of CA is the cell unit and the spatial structure created by a finite set of cells
- CA models usually are based on regular square cells that are obtained from remote sense maps (the cell is the pixel)
- The usual range of resolutions used on CA models varies from 500×500 square meters up to 25×25 square meters



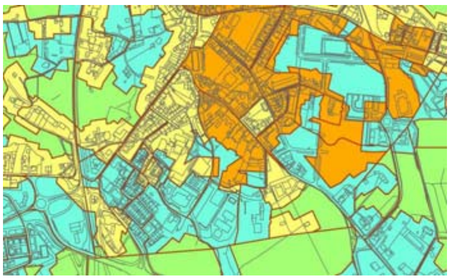
Methodology

Cells and Cell Space [2]

- Regular cells may not reproduce well the Portuguese urban structure. The consideration of regular square cells could became a problem, not only because of the resolution of the cell, but mainly because of data availability for this spatial partition
- Information is referenced to spatial census tracts that are drawn considering the urban structure
- If a high resolution was considered, the model would became closer to agent-based simulation than to CA with the cell's dimension becoming similar to the dimension of the land parcel → Data availability
- If the choice was for a low resolution of the cell, (100×100 m^2 or larger) the cell would become less representative of the urban form, particularly in the definition of urban areas → Spatial representativeness

Methodology

Cells and Cell Space [3]



Cell Space based on Census Tracts

Methodology

Cell States [1]

- The establishment of the set of land use states followed the main goal of keeping the model simple, as close to the CA formulation as possible
- Small urban areas are characterized by more simple urban structures and less diversity of land uses and functions
- Only six aggregate cell land uses were considered :
  - Urban Low Density (ULD)
  - Urban High Density (UHD)
  - Industrial (IND)
  - Non-urbanized urban areas (N-UUrb)
  - Non-urbanized industrial areas (N-UInd)
  - Restricted areas (Rest)

Methodology

Cell States [2]

- It was also considered the existence of a period of inactivity for a cell after it has suffered a change of state
- Urban areas were classified by their construction densities
- This classification is close to the planning rules commonly applied by municipal master plans, based on construction factors that are area dependent
- The consideration of all urban functions aggregated under a general urban land use incorporates those areas occupied by infrastructures, public facilities, and public space
- Cell suitability is assumed to be a binary variable that takes the value 1 if the cell is suitable for a given land use (both physically and legally) and 0 otherwise

Methodology

Cell States [3]

- Accessibility is assessed through the calibrated distance from the cell to the main functional centres of the territory (depending on the distance in time): the civil parish, the municipality main town and the main industrial area

$$AccMeas_i = \alpha_{acc} \times D_{i,MunicipalMainTown} + \beta_{acc} \times D_{i,CivilParish} + \gamma_{acc} \times D_{i,IndustrialArea}$$

- Road network was classified in three classes: main roads, secondary roads and local accesses. Each cell is served by at least a local access road

Methodology

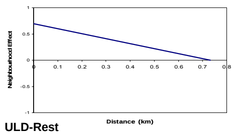
Cell Neighbourhood [1]

- One of the most distinctive characteristics of CA is its dependence on the interactions that are generated between each cell and its neighbours
- The first law of geography: everything is related to everything else, but near things are more related than distant things (Waldo Tobler)
- This relation is valid for land uses and for urban functions: two locations with similar or different land uses generates attraction or repulsion that varies with the distance between them
- The classical formulation considers that the neighbour effect is related only with the set of cells that are at a small distance to a given cell. The neighbourhood distance was considered as a model parameter

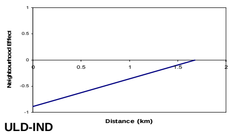
Methodology

Cell Neighbourhood [2]

- Neighbourhood effect was defined as a measure of the level of interaction between two locations within a certain distance that decays as the distance between land use location increases
- This value is set to -1 if two land uses manifest repulsion, 0 if they do not interact, and 1 if two land uses manifest attraction
- These relations are very difficult to assess, as they depend on several interdependent factors such as land value, housing demand, public facilities location among several others



ULD-Rest



ULD-IND

Methodology

Transition Rules [1]

- This process takes into account all the characteristics of each cell by the form of a measure of the transition potential, a calibrated value of accessibility, suitabilities and neighbourhood effect within the considered neighbourhood with a stochastic perturbation

$$Pot_{i,s}{}^{t} = f(Acc_i, Suit_{i,s}{}^{t}, NEff_{i,s}{}^{t}, \xi), i \in C; s \in CS$$

- $Pot_{i,s}{}^{t}$ is the transition potential for cell state s of cell i at time step t
- Acc_i is the accessibility measure for cell i (constant during the entire simulation period)
- $Suit_{i,s}{}^{t}$ is the suitability value for cell state s of cell i at time step t
- $NEff_{i,s}{}^{t}$ is the neighbourhood effect for cell state s of cell i at time step t
- ξ is the stochastic perturbation parameter

- The model is based on a population density demand approach. The model calculates the increase of population during the historical period from reference data and distributes that population over the territory under consideration

Methodology

Transition Rules [2]

- Four transition rules were considered. Cells can transit from
  - state 1 (ULD) to state 2 (UHD)
  - state 4 (N-UUrb) to states 1 and 2
  - state 5 (N-UInd) can transit to state 3 (Ind), with the criterion changing from population to employment
- Basically, one cell on state n can transit to state m after verifying a set of conditions:
  - The cell is not considered in the construction transitory condition
  - There still is population to distribute over the territory
  - Cell is at state n and its highest transition potential is associated to state m at time step t, chosen within the sub-set of cells that present the highest transition potential and are available for transition

Methodology

Measuring CA performance

- Contingency matrixes and associated *kappa* measures (*kValue*) were used to assess the quality of the simulation
- The contingency matrix compares the map resulting from the simulation with a reference map for the same year
- The consideration of the entire set of cell states for the calculation of the *kValue* value would produce a distortion on its significance as an indicator of a good simulation
- To avoid this distortion, a modification of the *kValue* measure was considered, called *ModkValue*
- This simplification consists in considering only the cell states that take part in the urban change dynamics, that is, not considering any cell state that is forced to remain in the same state throughout the simulation (and during real world evolution) or that does not participate in the land use dynamics

Methodology

The calibration process [1]

- Calibration is based on an optimization procedure that uses a new algorithm called Particle Swarm
- The choice for an optimization approach is related with the high number of parameters of the model
- Given such complexity – there are 38 calibration parameters that are highly interdependent – the use of sensitive analysis for each parameter would not reproduce the behaviour of one parameter when considering the others
- The main goal of an optimization approach is to ensure an extensive search for solutions in the multi-dimensional space through the minimization or maximization of an objective function related with the fitness function that assesses the model results

Methodology

The calibration process [2]

- The fitness function chosen to assess the quality of the model results was the modified *kappa* value, *ModkValue*
- The objective function is the maximization of *ModkValue*, which can assume the optimum value of 1 – absolute coincidence between simulation and reality
- Particle Swarm is based on a group of 60 particles that fly through the search space during a maximum of 100 iterations
- Each particle has 38 dimensions, one dimension per calibration parameter
- Each particle has a memory of its past search history, usually called the cognitive component, stored as the individual best position
- Each particle also "knows" the search results of the swarm, called the social component, stored as the best position of the entire swarm

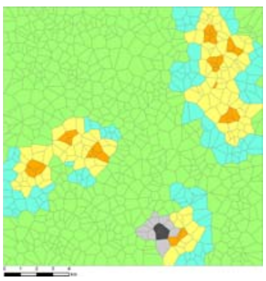
Test Problems

Definition

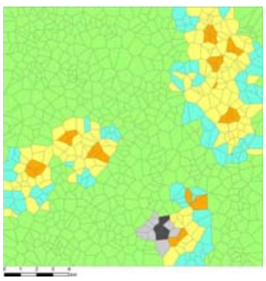
- A set of 20 theoretical problems was produced to test the performance of the CA model. The main goal is to evaluate the performance of the model for a group of theoretical problems, generated in a randomly manner under a set of conditions that could replicate likely real-world spatial structures
- The final goal of this procedure is to generate spatial structures that can be considered similar to an average small Portuguese municipality, both in its maximum dimension and in its number of cells
- These territories are square areas with their side dimension varying from 10 km to 20 km and were structured based on cells that intend to represent census spatial tracts. The number of cells varies from 800 to 1200

Test Problems

Problem 08 [1]



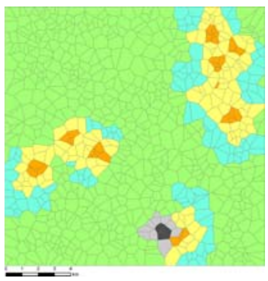
Initial Land Use Map



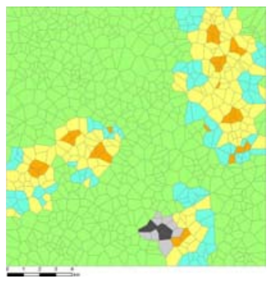
Final Land Use Map

Test Problems

Problem 08 [2]



Initial Land Use Map

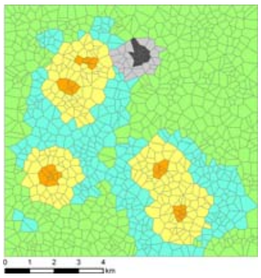


Simulation Land Use Map

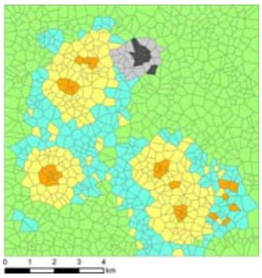
kValue=0.841

Test Problems

Problem 11 [1]



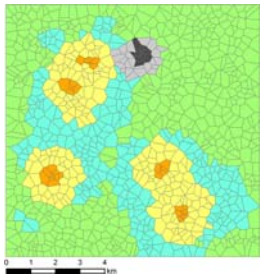
Initial Land Use Map



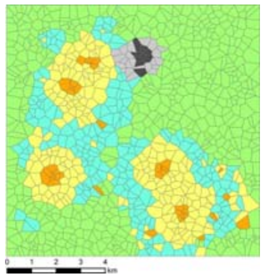
Final Land Use Map

Test Problems

Problem 11 [2]



Initial Land Use Map



Simulation Land Use Map

kValue=0.811

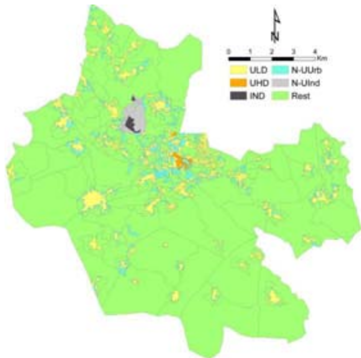
Case Study of Condeixa-a-Nova

Definition

- Condeixa-a-Nova registered a 48% growth rate for built area and a 17.8% growth rate for population for the period 1991/2001
- The cell space has 1433 cells resulting from the intersection of census tracts with urban areas for 1991 and 2001
- The set of cell states was identical to the one used for the test problems with 6 aggregate cell states
- Population densities were set to 6.6 and 13.2 inhabitants per hectare for state ULD and for state UHD in 2001, respectively
- The calibration procedure was based on a set of 60 particles with a maximum number of 100 iterations, for 10 years-long time steps

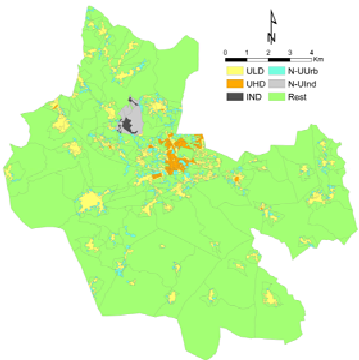
Case Study of Condeixa-a-Nova

Reference 1991



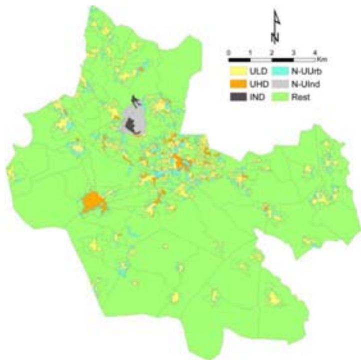
Case Study of Condeixa-a-Nova

Reference 2001



Case Study of Condeixa-a-Nova

Calibration 2001



kValue=0.621

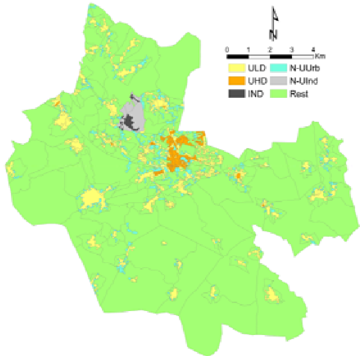
Case Study of Condeixa-a-Nova

Prospective analysis

- A prospective analysis was made after the calibration process considering the reference data from the national population censuses of 1991 and 2001
- This prospective analysis was produced with the set of calibration parameters obtained from the calibration phase
- Two prospective periods of 10 years were considered to obtain simulations for the years 2011 and 2021
- The population growth rate was considered constant of 10% in the two prospective periods

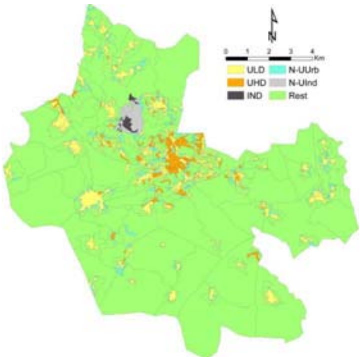
Case Study of Condeixa-a-Nova

Reference 2001



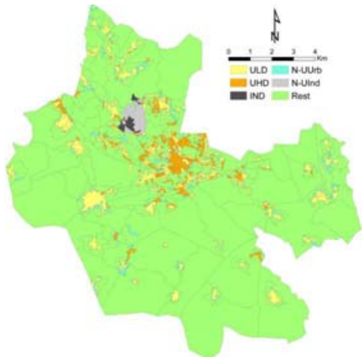
Case Study of Condeixa-a-Nova

Simulation 2011



Case Study of Condeixa-a-Nova

Simulation 2021



Conclusions

Model Results [1]

- The CA model presented promising results for the set of theoretical test problems, although a series of improvements must be implemented to overcome some shortcomings
- The assessed performance for this model is in line with the results of some other works published for equivalent measures
- Results must also be considered good because they are close to commonly accepted thresholds
- Future developments must focus the assessment of more sophisticate and accurate measures of accessibility, suitability and neighbourhood relations
- Neighbourhood must also be matter of further research

Conclusions

Model Results [2]

- The use of irregular cells based on census tracts is considered an important development of the approach. It shifts the focus of the model from land use – usually obtained from remote sensing image – to spatial information on both land use and demographics – census tracts contain disaggregate demographic and socio-economic information and can be easily classified in terms of land use
- Another innovative element of the study is the consideration of land use demand as a function of observed and expected population densities. It can be said that urban growth depends on population growth and on the variation of construction density